



## PATENT ABSTRACTS OF JAPAN

(11) Publication number: **11203403 A**

(43) Date of publication of application: 30.07.99

(51) Int. Cl.

**G06K 9/20****G06K 9/20****G06K 9/20**(21) Application number: **10006716**(71) Applicant: **OKI ELECTRIC IND CO LTD**(22) Date of filing: **16.01.98**(72) Inventor: **TORIGOE MAKOTO**

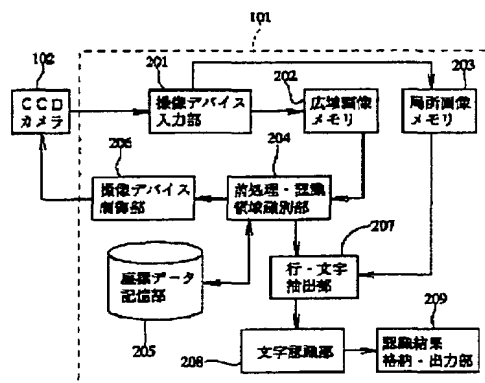
## (54) INFORMATION PROCESSOR

## (57) Abstract:

**PROBLEM TO BE SOLVED:** To provide a space-saving information processor requiring small reading time and reducing operator's working load.

**SOLUTION:** The whole original is photographed by a CCD camera 102 at low resolution, image data inputted from an image pickup device input part 201 are stored in a wide range picture memory 202 and a character recognition area is extracted by a preprocessing/recognition area identification part 204. Then a character recognition area is photographed by the camera 102 at high resolution under the control of an image pickup device control part 206, its picture data are stored in a local image memory 203, and after recognizing characters by a row/character extraction part 207 and a character recognition part 208, the recognized result is stored in a recognized result storing/outputting part 209.

COPYRIGHT: (C)1999,JPO



(51) Int.Cl.<sup>8</sup>

G 0 6 K 9/20

識別記号

3 4 0

F I

G 0 6 K 9/20

3 4 0 J

3 4 0 B

3 4 0 K

3 1 0 G

3 2 0 F

3 1 0

3 2 0

審査請求 未請求 請求項の数 5 O L (全 14 頁)

(21) 出願番号

特願平10-6716

(22) 出願日

平成10年(1998) 1月16日

(71) 出願人 000000295

沖電気工業株式会社

東京都港区虎ノ門1丁目7番12号

(72) 発明者 鳥越 真

東京都港区虎ノ門1丁目7番12号 沖電気  
工業株式会社内

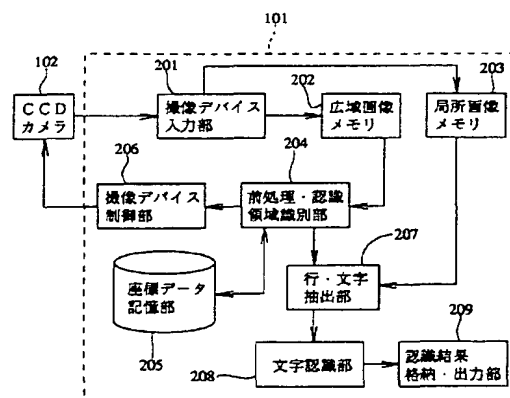
(74) 代理人 弁理士 大垣 孝

(54) 【発明の名称】 情報処理装置

(57) 【要約】

【課題】 省スペース型で、読み取り時間が速く、オペレータの作業負担も小さい情報処理装置を提供する。

【解決手段】 まず、CCDカメラ102で原稿全体を低解像度で撮像して、画像データを撮像デバイス入力部201から広域画像メモリ202に格納し、前処理・認識領域識別部204で文字認識領域を抽出する。そして、撮像デバイス制御部206の制御によりCCDカメラ102で文字認識領域を高解像度で撮像して局所画像メモリ203に画像データを格納し、行・文字抽出部207および文字認識部208で文字認識を行った後、この結果を認識結果格納・出力部209に格納する。



第1の実施の形態の構成図

## 【特許請求の範囲】

【請求項 1】 原稿を二次元的に読み取る撮像デバイスと、  
この撮像デバイスが取得した画像データから文字認識実行領域を抽出した後、この文字認識事項領域に対応する画像データから行および文字を抽出し、抽出された文字の文字認識を行う本体部と、  
を備えたことを特徴とする情報処理装置。

【請求項 2】 前記本体部が、  
前記撮像デバイスが取得した画像データを入力する撮像デバイス入力部と、  
前記撮像デバイス入力部が低解像度で入力した前記原稿全面の画像データを一時的に記憶する広域画像メモリと、  
この広域画像メモリから読み出した前記原稿全面の画像データから文字認識実行領域を識別する認識領域識別部と、  
この認識領域識別部が識別した前記文字認識実行領域を高解像度で撮像するように前記撮像デバイスを制御する撮像デバイス制御部と、  
前記撮像デバイス入力部が前記撮像デバイスから高解像度で入力した前記文字認識実行領域の画像データを一時的に記憶する局所画像メモリと、  
この局所画像メモリから読み出した前記文字認識実行領域の画像データから行および文字を抽出する行・文字抽出部と、  
この行・文字認識部が抽出した文字の文字認識を行う文字認識部と、  
を備えたことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 3】 前記本体部が、  
前記撮像デバイスが取得した画像データを入力する撮像デバイス入力部と、  
前記撮像デバイス入力部が低解像度で入力した前記原稿全面の画像データを一時的に記憶する広域画像メモリと、  
この広域画像メモリから読み出した前記原稿全面の画像データから文字認識実行領域を識別する認識領域識別部と、  
前記撮像デバイス入力部が入力した前記原稿全面の画像データを所定時間ごとに取り込んで一時的に記憶するスナップショットメモリと、  
このスナップショットメモリから順次読み出した画像データから縦横方向の黒ランがともに所定数を越える黒ラン領域をそれぞれ検出し、この黒ラン領域の位置変化と前記認識領域識別部の識別結果とを用いてオペレータが指示する前記文字認識実行領域を検出する指示先検出部と、  
この指示先検出部が検出した前記文字認識実行領域を高解像度で撮像するように前記撮像デバイスを制御する撮

像デバイス制御部と、

前記撮像デバイス入力部が前記撮像デバイスから高解像度で入力した前記文字認識実行領域の画像データを一時的に記憶する局所画像メモリと、  
この局所画像メモリから読み出した前記文字認識実行領域の画像データから行および文字を抽出する行・文字抽出部と、  
この行・文字認識部が抽出した文字の文字認識を行う文字認識部と、  
を備えたことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 4】 前記本体部が、  
前記撮像デバイスが取得した画像データを入力する撮像デバイス入力部と、  
前記撮像デバイス入力部が低解像度で入力した前記原稿全面の画像データを一時的に記憶する広域画像メモリと、  
前記撮像デバイス入力部が入力した前記原稿全面の画像データを所定時間ごとに取り込んで一時的に記憶するスナップショットメモリと、  
このスナップショットメモリから順次読み出した画像データと前記広域画像メモリから読み出した画像データとの差分を順次求め、この差分の変化を用いてオペレータが文字認識の実行を指示する指示領域を検出する差分検出部と、  
この差分検出部が検出した前記指示領域を高解像度で撮像するように前記撮像デバイスを制御する撮像デバイス制御部と、  
前記撮像デバイス入力部が前記撮像デバイスから高解像度で入力した前記指示領域の画像データを一時的に記憶する局所画像メモリと、  
この局所画像メモリから読み出した前記文字認識実行領域の画像データから行および文字を抽出する行・文字抽出部と、  
この行・文字認識部が抽出した文字の文字認識を行う文字認識部と、  
を備えたことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 5】 前記本体部が、  
前記撮像デバイスが取得した画像データを入力する撮像デバイス入力部と、  
前記撮像デバイス入力部が入力した前記原稿全面の画像データを所定時間ごとに取り込んで一時的に記憶するスナップショットメモリと、  
このスナップショットメモリから読み出した画像データから、所定長さ以上の白ランで囲まれた領域を抽出し、この領域が 1 箇所のみであるときに、この領域をオペレータが文字認識の実行を指示する前記指示領域であると認識する認識領域識別部と、  
前記撮像デバイスを用いて前記撮像デバイス入力部が取

得した、前記文字認識実行領域の画像データから、行および文字を抽出する行・文字抽出部と、この行・文字認識部が抽出した文字を認識する文字認識部と、を備えたことを特徴とする請求項1に記載の情報処理装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、スキャナで読み取った文書画像から任意の文字情報領域を抽出して文字認識を行う情報処理装置に関するものである。

【0002】

【従来の技術】従来より、スキャナで原稿を読み取り、これによって得られた画像データから任意の文字情報領域を抽出して文字認識を行う情報処理装置が知られている。

【0003】かかる情報処理装置で使用されるスキャナとしては、従来、フラットベット型のもものとハンディ型のものとのが知られている。

【0004】フラットベット型スキャナでは、原稿載置台に原稿を密着させた状態で、スキャナヘッドまたは原稿載置台を移動させることにより、原稿の読み取りを行う。また、ハンディ型スキャナでは、原稿にスキャナヘッドを密着させて手動で移動させることにより、原稿の読み取りを行う。これらのスキャナは、いずれもスキャナヘッドに一次元イメージセンサが設けられている。

【0005】また、フラットベット型スキャナを用いた情報処理装置として、原稿をいきなり高解像度で読み取るのではなく、まず低解像度でプリスキャンを行って全体レイアウトをディスプレイ装置に表示させ、このディスプレイ装置等を用いてオペレータが原稿上の任意の領域を指定した後、かかる指定領域のみを高解像度で本スキャンして再度読み取る形式のものも、従来知られている。

【0006】

【発明が解決しようとする課題】しかしながら、従来の情報処理装置には、以下のような欠点があった。

【0007】フラットベット型スキャナを用いた情報処理装置には、このフラットベット型スキャナのサイズが大きいので設置スペースを広く取る必要があるという欠点があった。

【0008】また、このフラットベット型スキャナを用いた情報処理装置には、例えば本や雑誌等、綴じ合わせた原稿を読み取る場合に、使い勝手が悪いという欠点もあった。すなわち、本や雑誌等は単に載置しただけでは読み取り面を原稿載置台に密着させることができず、また、載置したときの安定性が悪いので、スキャナでの読み取り時に読み取り面が180度開くようにオペレータが手で押さえつける必要がある。このため、一人での作業が困難な場合が多く、また、本や雑誌等の綴じ合わせ

部分を破損させてしまう場合もある。

【0009】さらに、かかる情報処理装置では、イメージセンサとして一次元のものを使用しており、スキャナヘッドまたは原稿載置台を移動させて原稿を読み取る構造となっているため、原稿の読み取りに時間がかかり、特にプリスキャンと本スキャンとを行う場合には読み取り時間が非常に長くなってしまいうという欠点があった。

【0010】加えて、かかる情報処理装置で全体をプリスキャンした後で一部の領域のみを本スキャンしようとする場合、副走査方向についてはスキャナヘッドまたは原稿載置台の移動を制御することによって読み取り部分を当該領域のみに限定することができるが、主走査方向については読み取り部分を限定することができない。このため、従来の情報処理装置では、主走査方向については、読み込むべき領域以外についても、一旦読み込むのと同様のスキャナヘッドまたは原稿載置台の移動を行う必要があり、この分だけ余分な動作を行うという欠点があった。

【0011】一方、ハンディ型スキャナを用いた情報処理装置では、サイズは小さいものの、上述のようにオペレータがスキャナヘッドを原稿に密着させて手動で移動させなければならず、また、このときの移動速度をほぼ一定にしないと高画質の画像データを得ることができないので、入力作業時のオペレータの負担が大きいという欠点があった。

【0012】このような理由から、省スペース型で、読み取り時間が速く、オペレータの作業負担も小さい情報処理装置の登場が囑望されていた。

【0013】

【課題を解決するための手段】この発明にかかる情報処理装置は、原稿を二次元的に読み取る撮像デバイスと、この撮像デバイスが取得した画像データから文字認識実行領域を抽出した後、この文字認識事項領域に対応する画像データから行および文字を抽出し、抽出された文字の文字認識を行う本体部とを備える。

【0014】このような構成によれば、二次元の撮像デバイスを用いたので、スキャナヘッドを移動させることなく原稿の読み取りを行うことができ、これにより、読み取り時間の短縮とオペレータの作業負担の軽減とを図ることができる。さらに、スキャナのサイズを小型化して、省スペース化を図ることができる。

【0015】

【発明の実施の形態】以下、この発明の実施の形態について、図面を用いて説明する。なお、図中、各構成成分の大きさ、形状および配置関係は、この発明が理解できる程度に概略的に示してあるにすぎず、また、以下に説明する数値的条件は単なる例示にすぎないことを理解されたい。

【0016】第1の実施の形態

まず、この発明の第1の実施の形態に係る情報処理装置

について、図1～図4を用いて説明する。

【0017】図1(A)、(B)は、この実施の形態に係る情報処理装置の全体構成を概念的に示す外観図である。

【0018】同図において、本体部101は、図2(後述)に示すような回路を備えている。

【0019】撮像デバイスとしてのCCD(Charge Coupled Device)カメラ102は、原稿110を読み取るためのスキャナとして使用される。このCCDカメラ102は、図1(A)に示したようにディスプレイ装置103(後述)に固定することとしてもよいし、図1(B)に示したように固定せずにオペレータが手に持って読み取りを行うようにしてもよい。さらに、このCCDカメラ102を取り外し自在にディスプレイ装置に装着して、オペレータの判断により、装着した状態または手で持った状態で使用できることとしてもよい。

【0020】ディスプレイ装置103は、CCDカメラ102で原稿を読み取ることによって得られた画像データ等を表示する。

【0021】キーボード104は、例えばCCDカメラ102に原稿を読み取らせる場合等、オペレータが情報処理装置を操作する場合に使用される。

【0022】図2は、この実施の形態の情報処理装置に係る回路構成の要部を示すブロック図である。

【0023】同図において、撮像デバイス入力部201は、CCDカメラ102が撮像した画像データを取り込んで、広域画像メモリ202または局所画像メモリ203に転送する。

【0024】広域画像メモリ202は、CCDカメラ102によって原稿全面が低解像度で撮像された場合に、この画像データを一時的に記憶する。

【0025】局所画像メモリ203は、CCDカメラ102によって原稿の文字認識実行領域が高解像度で撮像されたときに、この画像データを一時的に記憶する。

【0026】前処理・認識領域識別部204は、広域画像メモリ202から原稿全面の画像データを取り込んで文字認識を行う領域(文字認識実行領域)を判別し、この文字認識実行領域の座標を検出する。

【0027】座標データ記憶部205には、例えば原稿用紙の枠構成や所定書類の書式等のデータが、座標データとして記憶されている。

【0028】撮像デバイス制御部206は、前処理・認識領域識別部204が判別した文字認識決定領域を高解像度で撮像するように、CCDカメラ102を制御する。

【0029】行・文字抽出部207は、局所画像メモリ203に記憶された画像データから、文字認識実行領域の文字および行の画像データを抽出する。

【0030】文字認識部208は、行・文字抽出部207が抽出した文字を取り込んで、文字認識を行う。

【0031】認識結果格納・出力部209は、文字認識部208が認識した文字をディスプレイ装置103(図1参照)に表示させたり、図示しないメモリに格納したりする。

【0032】なお、図2では広域画像メモリ202と局所画像メモリ203とを別個に設けることとしたが、1個のメモリを共用することとしてもよい。

【0033】次に、この実施の形態に係る情報処理装置の動作について、図3および図4を用いて説明する。

【0034】まず、CCD102が原稿全体のプリスキャンを行うと、これによって得られた画像データ(図4(A)参照)を撮像デバイス入力部201が取り込み(S301)、広域画像メモリ202に格納する(S302)。このときの解像度は、前処理・認識領域識別部204が文字認識実行領域を判断できる程度であればよいので、例えば50dpi(dot per inch)程度の低い解像度で十分である。

【0035】次に、前処理・認識領域識別部204が、この画像データを広域画像メモリ202から読み出して、文字認識実行領域G11、G12、G13(図4(A)参照)を抽出する(S303)。ここで、文字認識実行領域G11、G12、G13を抽出する方法としては、例えば原稿全体に相当する画像データを走査し、黒ランおよび白ランを検出して結合する方法があるが、公知の技術であるので説明を省略する。また、このときに、画像データの画質が悪いこと等の原因によって文字認識実行領域G11、G12、G13の抽出ができないときや、フォーマットが予め明らかであるような場合或いは抽出した文字認識実行領域G11、G12、G13によって文書フォーマットを特定することが可能な場合には、座標データ記憶部205に記憶された座標データを参照して文字認識実行領域G11、G12、G13の抽出が行われる。

【0036】続いて、撮像デバイス制御部206が、前処理・認識領域識別部204から文字認識実行領域G11の座標を入力し、この文字認識実行領域の撮像(すなわち本スキャン)を行うようにCCDカメラ102を制御する(S304)。このときの解像度は、文字認識部208が文字認識を行うことができる解像度であることが必要であり、例えば200dpi以上とすることが望ましい。ここで、文字認識実行領域が複数箇所あると前処理・認識領域識別部204が判断した場合には、最初の文字認識実行領域G11のみが、CCDカメラ102によって撮像される。また、局所画像メモリ203の容量の制限やCCDカメラ102の解像度の制限により文字認識実行領域G11全体を一回の撮像動作で撮像することができない場合には、この局所画像メモリ203が記憶できる部分或いはCCDカメラ102が撮像可能な部分のみを段落単位で選択し、この部分の撮像を、一回目の撮像として行う(図4(B)参照)。

【0037】そして、この撮像によって得られた画像デ

ータが、撮像デバイス入力部201に取り込まれ、局所画像メモリ203に格納される(S305)。

【0038】次に、行・文字抽出部207が局所画像メモリ204に記憶された文字認識実行領域の画像データから文字および行の画像データを抽出する(S306)。そして、行・文字抽出部207が抽出した文字を用いて、文字認識部208が文字認識を行う(S307)。

【0039】その後、文字認識実行領域G<sub>11</sub>のすべての領域についての文字認識が終了したか否かを判断する

(S308)。そして、文字認識実行領域G<sub>11</sub>に文字認識されていない部分があると判断された場合には、撮像デバイス制御部206がCCDカメラ102の撮像位置をシフトさせ(S309)、次の撮像を行った後(S304)、ステップS305～S308を繰り返す。この動作は、ステップS308において文字認識実行領域G<sub>11</sub>全体についての文字認識が終了したと判断されるまで繰り返される。

【0040】一方、ステップS308で文字認識実行領域G<sub>11</sub>全体についての文字認識が終了したと判断された場合には、続いて、すべての文字認識実行領域G<sub>11</sub>、G<sub>12</sub>、G<sub>13</sub>についての文字認識が終了したか否かを判断する(S310)。そして、文字認識が行われていない文字認識実行領域(ここではG<sub>12</sub>、G<sub>13</sub>)が存在すると判断された場合は、撮像デバイス制御部206がCCDカメラ102の撮像位置を移動させ(S311)、次の撮像(ここでは文字認識実行領域G<sub>12</sub>、図4(C)参照)を行った後(S304)、ステップS305～S310を繰り返す。これら一連の動作は、ステップS310においてすべての文字認識実行領域G<sub>11</sub>、G<sub>12</sub>、G<sub>13</sub>全体についての文字認識が終了したと判断されるまで繰り返される。

【0041】ステップS310で、すべての文字認識実行領域G<sub>11</sub>、G<sub>12</sub>、G<sub>13</sub>全体についての文字認識が終了したと判断された場合には、文字認識部208の認識結果が認識結果格納・出力部209に送られる(S312)。これにより、文字認識部208によって認識された文字がディスプレイ装置103(図1参照)に表示されるなどして、情報処理装置による処理が終了する。

【0042】このように、この実施の形態に係る情報処理装置によれば、CCDカメラ102をスキャナとして用いることとしたので、原稿の読み取り時にスキャナヘッドまたは原稿載置台を移動させる必要が無く、したがって処理速度を向上させることができる。

【0043】また、本スキャンを行う際に、主走査方向・副走査方向ともに任意の範囲を選択して撮像することができる(すなわち文字認識実行領域のみに限定して撮像を行うことができるので)、必要なメモリ容量を低減することができる。また、上述したように広域画像メモリ202と局所画像メモリ203とを1個のメモリの

共用にした場合には、さらにメモリ容量を低減することが可能である。

【0044】また、プリスキャンと本スキャンとで解像度が異なるため、CCD102等の撮像デバイスとして比較的低下解像度のものを使用することができ、低価格化を図る上で有効である。

【0045】さらに、フラットベット型スキャナを用いる場合と比較して設置スペースを小さくすることができる。

【0046】加えて、この実施の形態に係る情報処理装置では、本や雑誌等、綴じ合わせた原稿を読み取る場合に、この原稿を180度開く必要が無く、90度以上に開いて片方のページずつ読み取らせればよい。このため、読み取り時にオペレータが原稿を押さえておく必要がない場合も多く、また、綴じ合わせ部分を破損するおそれもない。

【0047】第2の実施の形態

次に、この発明の第2の実施の形態について、図5～図8を用いて説明する。

【0048】図5は、この実施の形態に係る情報処理装置の全体構成を概念的に示す外観図である。

【0049】同図において、図1と同じ符号を付した構成部は、それぞれ図1の場合と同じものを示している。

【0050】この実施の形態では、図1(A)に示した場合と同様、撮像デバイスとしてのCCDカメラ102は、ディスプレイ装置103に固定されている。

【0051】また、後述するように、本体部101は、オペレータが原稿110の文字認識実行領域を指やペン等で指し示して指定したときに、この指定がなされた文字認識実行領域のみについて文字認識を行うように構成されている。

【0052】図6は、この実施の形態の情報処理装置に係る回路構成の要部を示すブロック図である。

【0053】同図において、図2と同じ符号を付した構成部は、それぞれ図2の場合と同じものを示している。

【0054】図6に示した情報処理装置は、指定領域検出部601を備えている点で、図2の場合と異なる。この指定領域検出部601は、スナップショットメモリ602と、指示先検出部603とを備えている。

【0055】スナップショットメモリ602は、後述するようにしてCCDカメラ102が所定時間ごとに原稿を撮像することによって得られた画像データを、順次入力して記憶する。

【0056】指示先検出部603は、スナップショットメモリ602に記憶された画像データを取り込んで縦方向および横方向についてそれぞれ黒ランを検出する。そして、縦方向の黒ラン値および横方向の黒ラン値が所定のしきい値以上であれば、これらの黒ランが指先またはペン先等によるものであると判断し、かかる指先等へ外接する矩形領域の左上頂点の座標を検出する。そして、

後述するようにして、左上頂点座標の最大値を求め、この最大値からオペレータが指示している文字認識実行領域を判断する。

【0057】なお、図6では広域画像メモリ202と局所画像メモリ203とスナップショットメモリ602とを別個に設けることとしたが、1個のメモリを共用することとしてもよい。

【0058】次に、この実施の形態に係る情報処理装置の動作について、図7～図9を用いて説明する。

【0059】まず、上述の第1の実施の形態と場合と同様、CCDカメラ102が低解像度で原稿全体のプリスキャンを行うと、これによって得られた画像データ(図9(A)参照)を撮像デバイス入力部201が取り込み(S701)、広域画像メモリ202に格納する(S702)。

【0060】さらに、前処理・認識領域識別部204が、この画像データを広域画像メモリ202から読み出し、上述の第1の実施の形態と同様にして、文字認識実行領域G21、G22、G23(図9(A)参照)を抽出する(S703)。このときに、画像データの画質が悪いこと等の原因によって文字認識実行領域G21、G22、G23の抽出ができないときや、フォーマットが予め明らかであるような場合或いは抽出した文字認識実行領域G21、G22、G23によって文書フォーマットを特定することが可能な場合には、座標データ記憶部205に記憶された座標データを参照して文字認識実行領域G21、G22、G23の抽出が行われる。

【0061】次に、CCDカメラ102が、原稿全体の撮像を低解像度で行い(S704)、この撮像によって得られた画像データをスナップショットメモリ602に格納する(S705)。続いて、指示先検出部603が、このスナップショットメモリ602から画像データを取り込んで、縦方向および横方向についてそれぞれ黒ランを検出する(S706)。そして、所定のしきい値以上の黒ランの有無を縦方向および横方向について判断し(S707)、黒ランが無い場合は、左上座標の最大値がすでに検出されているか否かの判断を行う(S708)。さらに、左上座標の最大値が検出されていない場合は、オペレータが原稿の文字認識実行領域を指やペン等で指し示す前であると判断して、ステップS704に戻り、このステップS704～S708を繰り返す。

【0062】ここで、オペレータが原稿の文字認識実行領域を指やペン等で指し示そうとした場合には、指やペン等が原稿の上を移動するので、ステップS707で黒ランが検出される(図9(B)参照)。この場合には、指示先検出部603が、この黒ラン領域のみを画像データ中から抽出した後(図9(C)参照)、この黒ラン領域に外接する矩形領域の左上頂点の座標(x, y)を検出する(S709)。そして、ステップS704に戻り、このステップS704～S708を繰り返す。これ

により、オペレータの指やペン等が原稿の上に位置している間は、所定の時間間隔ごとにCCDカメラ102による撮像と左上頂点座標の検出とが繰り返されることとなる。そして、検出された左上頂点座標の最大値(原稿の右下端部(x0, y0)からの距離が最も長いもの)を随時検出する。

【0063】オペレータが原稿の文字認識実行領域を右手の指や右手に持ったペン等で指し示す場合、オペレータの手は、原稿の右下端部付近から左上端部方向に移動し、指やペン等の先端部で所望の文字認識実行領域付近を指し示した後で、右下方向に戻る。そして、このときの左上端部座標(x, y)は、指先やペン先等の位置に相当する。従って、黒ラン領域の位置検出を所定時間間隔で繰り返し、左上頂点座標(x, y)の最大値を検出することにより、オペレータが指やペン等で指し示した位置を検出することが可能となる。

【0064】オペレータが原稿の文字認識実行領域の指示を終了し、指やペン等が原稿外側に出ると、ステップS707で黒ランが検出されなくなる。この場合には、ステップS708で左上座標の最大値が検出されていると判断されるので、続いて、この最大左上座標(xm, ym)と文字認識実行領域G21、G22、G23の座標とを比較する(S709)。そして、最大左上座標(xm, ym)が、文字認識実行領域G21、G22、G23のいずれかの領域内に含まれているときは、その文字認識実行領域がオペレータによって指定されたと判断して、後述のステップS712以降を実行する。例えば、図9(D)に示したように、最大左上座標(xm, ym)が文字認識実行領域G23に含まれるときは、オペレータが文字認識実行領域G23を指し示したものと判断する。

【0065】ステップS709で、最大左上座標(xm, ym)が、文字認識実行領域G21、G22、G23のいずれにも含まれていない場合は、各文字認識実行領域G21、G22、G23の左上付近に走査領域S21、S22、S23を設定し(図9(E)参照)、これらの領域S21、S22、S23を最大左上座標(xm, ym)と比較する(S710)。そして、最大左上座標(xm, ym)が領域S21、S22、S23のいずれかに含まれている場合は、その文字認識実行領域がオペレータによって指定されたと判断して、後述のステップS712以降を実行する。

【0066】ステップS710で、最大左上座標(xm, ym)が、領域S21、S22、S23のいずれにも含まれていない場合は、最大左上座標(xm, ym)の左上方向にある文字認識実行領域の内でも最大左上座標(xm, ym)から近い領域を検索する(S711)。そして、この領域がオペレータによって指定されたと判断して、ステップS712以降を実行する。

【0067】ステップS712では、撮像デバイス制御部206が、オペレータが指定した文字認識実行領域の

座標を前処理・認識領域識別部 204 から入力し、この文字認識実行領域の撮像（すなわち本スキャン）を行うように CCD カメラ 102 を制御する。このときの解像度は、文字認識部 208 が文字認識を行うことができる解像度であることが必要であり、例えば 200 dpi 以上とすることが望ましい。ここで、局所画像メモリ 203 の容量の制限や CCD カメラ 102 の解像度の制限により文字認識実行領域全体を一回の撮像動作で撮像することができない場合には、この局所画像メモリ 203 の容量分の撮像或いは CCD カメラ 102 が撮像可能な部分のみを、一回目の撮像として行う。

【0068】そして、撮像デバイス制御部 206 の制御によって CCD カメラ 102 が本スキャンを行うと、この撮像によって得られた画像データが、撮像デバイス入力部 201 に取り込まれ、局所画像メモリ 203 に格納される（S713）。

【0069】次に、行・文字抽出部 207 が局所画像メモリ 203 に記憶された文字認識実行領域の画像データから文字および行の画像データを抽出する（S714）。そして、行・文字抽出部 207 が抽出した文字を用いて、文字認識部 208 が文字認識を行う（S715）。

【0070】その後、オペレータによって指定された文字認識実行領域全体についての文字認識が終了したか否かを判断する（S716）。そして、文字認識実行領域に文字認識されていない部分があると判断された場合には、撮像デバイス制御部 206 が CCD カメラ 102 の撮像位置をシフトさせ（S717）、次の撮像を行った後（S712）、ステップ S713～S716 を繰り返す。この動作は、ステップ S716 において文字認識実行領域全体についての文字認識が終了したと判断されるまで繰り返される。

【0071】ステップ S716 で、文字認識実行領域全体についての文字認識が終了したと判断された場合には、文字認識部 208 の認識結果が認識結果格納・出力部に送られる（S718）。これにより、文字認識部 208 によって認識された文字がディスプレイ装置 103（図 5 参照）に表示される等して、情報処理装置による処理が終了する。

【0072】なお、文字認識実行領域 G<sub>21</sub>、G<sub>22</sub>、G<sub>23</sub> の内の 2 箇所以上を指定できるようにしたい場合には、ステップ S704～S717 を繰り返し実行するように、画像処理装置を構成すればよい。

【0073】また、文字認識実行領域の指定が正常に終了したことを、ピープ音やランプ等によって、オペレータに告知することとしてもよい。

【0074】このように、この実施の形態に係る画像処理装置によれば、文字認識を行いたい領域をオペレータが指定する際に、原稿を指やペン等で直接指し示すだけでよいので、従来のようにディスプレイ装置の画面上で

指定する場合と比較して、装置の使い勝手を向上させることができる。

【0075】また、処理速度を向上させることができる点、メモリ容量を低減することができる点、撮像素子（ここでは CCD カメラ 102）として比較的低解像度のものを使用することができる点、設置スペースを小さくすることができる点および本や雑誌等をオペレータが原稿を押さえる必要や綴じ合わせ部分を破損するおそれがない点は、第 1 の実施の形態と同様である。

#### 【0076】第 3 の実施の形態

次に、この発明の第 3 の実施の形態について、図 10～図 13 を用いて説明する。

【0077】ここで、この実施の形態では、図 1（A）に示した場合と同様、撮像デバイスとしての CCD カメラ 102 は、ディスプレイ装置 103 に固定されている。

【0078】また、後述するように、本体部 101 は、オペレータが原稿の所望の領域の左上端部および右下端部を指やペン等で指し示して指定したときに、この指定がなされた領域のみについて文字認識を行うように構成されている。

【0079】図 10 は、この実施の形態の情報処理装置に係る回路構成の要部を示すブロック図である。

【0080】同図において、図 2 と同じ符号を付した構成部は、それぞれ図 2 の場合と同じものを示している。

【0081】図 10 に示した情報処理装置は、指定領域検出部 1001 を備えている点で、図 2 の場合と異なる。この指定領域検出部 1001 は、スナップショットメモリ 1002 と、差分検出部 1003 とを備えている。

【0082】スナップショットメモリ 1002 は、上述の第 2 の実施の形態の場合と同様、CCD カメラ 102 が所定時間ごとに原稿を撮像することによって得られた画像データを、順次入力して記憶する。

【0083】差分検出部 1003 は、後述するようにして、指やペン等を認識する動作を 2 回行う。そして、認識された指やペン先等の外接矩形の 2 種類の左上頂点座標から特定される領域を、オペレータが指定した文字認識実行領域であると判断して、この判断結果を行・文字抽出部 207 に出力する。

【0084】なお、図 10 では広域画像メモリ 202 と局所画像メモリ 203 とスナップショットメモリ 1002 を別個に設けることとしたが、1 個のメモリを共用することとしてもよい。

【0085】次に、この実施の形態に係る情報処理装置の動作について、図 11～図 13 を用いて説明する。

【0086】まず、上述の第 1 の実施の形態と同様、CCD カメラ 102 が低解像度で原稿全体のプリスキャンを行うと、これによって得られた画像データ（図 13（A）参照）を撮像デバイス入力部 201 が取り込



み(S1101)、広域画像メモリ202に格納する(S1102)。

【0087】次に、CCDカメラ102が、原稿全体の撮像を行い(S1103)、この撮像によって得られた画像データをスナップショットメモリ1002に格納する(S1104)。このときの解像度は、上述のステップS1101と同じであることが望ましい。続いて、差分検出部1003が、スナップショットメモリ1002を走査して、このスナップショットメモリ1002に格納された画像データと広域画像メモリ202に格納された画像データとの差分(すなわち両画像データが異なる領域)を検出する(S1105)。さらに、差分が検出されたか否かを判断し(S1106)、差分が無い場合は、始点座標および終点座標(後述)がすでに検出されているか否かの判断を行う(S1107)。ここで、始点座標および終点座標が検出されていない場合は、原稿の文字認識を行いたい領域を指やペン等で指し示す動作をオペレータが開始する前であると判断して、ステップS1103に戻り、各ステップS1103~S1107を繰り返す。

【0088】ここで、オペレータが原稿の文字認識実行領域を指やペン等で指し示そうとした場合には、原稿上に指やペン等が存在するので、この指やペン等の部分の画像データは撮像デバイス入力部201がステップS1101で入力した画像データと異なる内容となり、この差異がステップS1106で「差分」として検出される

(図13(B)参照)。この場合には、差分検出部1003が、この差分領域のみを画像データ中から抽出した後(図13(C)参照)、この差分領域に外接する矩形領域の左上頂点の座標(x, y)を検出する(S1108)。そして、このときに検出された左上頂点座標

(x, y)が、それまでに検出された左上頂点座標の最大値(以下「最大座標」と記す)より大きいのか否かの判断を行う(S1109)。ここで、最大座標よりも小さかった場合は、この左上頂点座標(x, y)が前回検出した左上頂点座標(x, y)と同じ座標であるか否かの判断を行う(S1110)。左上頂点座標(x, y)が前回検出したものと同じ座標でなかった場合は、ステップS1103に戻り、このステップS1103~S1110を繰り返す。すなわち、情報処理装置は、所定の時間間隔ごとにCCDカメラ102による撮像と左上頂点座標の検出とを繰り返し、検出された左上頂点座標(x, y)のうちで最も大きいもの(原稿の右下端部(x<sub>0</sub>, y<sub>0</sub>)からの距離が最も長いもの)を、随時、最大座標として記憶する。

【0089】この実施の形態では、原稿の文字認識実行領域を右手の指や右手に持ったペン等で指し示す場合、オペレータは、まず、原稿の右下端部付近から左上端部方向に指やペン等を移動させて、文字認識を行いたい部分の左上端部を指し示し、次に、指やペン等を右下方向

に移動させて、文字認識を行いたい部分の右下端部を指し示す(このとき、指やペン等の動作は一旦停止する)。また、上述の左上端部座標は、指先やペン先等の位置に相当する。

【0090】従って、オペレータが文字認識を行いたい部分の左上端部を指し示すために指やペンを左上方向に移動させているときには、ステップS1109が実行される度に最大座標が更新される。この座標は、始点座標として差分検出部1103に一時的に格納される(S1111)。その後、ステップS1103以降が実行される。

【0091】次に、オペレータが文字認識を行いたい部分の右下端部を指し示すために指やペンを右下方向に移動させると、このときには、ステップS1108で検出された差分外接矩形の左上座標は最大座標よりも小さくなる。但し、オペレータが当該部分の右下端部を指し示したときには、指やペン等の動作は一旦停止するので、差分外接矩形の左上座標は前回検出された左上座標と一致する(図13(D)参照)。すなわち、ステップS1108で検出された差分外接矩形の左上座標が最大座標と一致せず(S1109)且つ前回検出した座標と一致した場合には(S1110)、この座標は、オペレータが文字認識を行いたい部分の右下座標に相当する。この座標は、終点座標として差分検出部に一時的に格納される(S1112)。その後、ステップS1103以降が実行される。

【0092】その後、オペレータが指やペン等を原稿上から離すと、CCDカメラ102が撮像する画像データは原稿のみになる。このため、ステップS1106では、差分画像がないと判断される。

【0093】続いて、ステップS1107では、始点座標および終点座標がすでに検出されていると判断される。これにより、情報処理装置は、原稿の文字認識を行いたい領域を指やペン等で指し示す動作をオペレータが終了した後であると判断して、ステップS1113以降を実行する。

【0094】ステップS1113では、撮像デバイス制御部206が、オペレータが指定した部分の座標を差分検出部1003から入力し、この指定領域の撮像(すなわち本スキャン)を行うようにCCDカメラ102を制御する。このときの解像度は、文字認識部208が文字認識を行うことができる解像度であることが必要であり、例えば200dpi以上とすることが望ましい。ここで、局所画像メモリ203の容量の制限やCCDカメラ102の解像度の制限により指定領域全体を一回の撮像動作で撮像することができない場合には、この局所画像メモリ203の容量分の撮像或いはCCDカメラ102が撮像可能な部分のみを、一回目の撮像として行う。

【0095】そして、撮像デバイス制御部206の制御によってCCDカメラ102が本スキャンを行うと、こ

10

20

30

40

50

の撮像によって得られた画像データが、撮像デバイス入力部201に取り込まれ、局所画像メモリ203に格納される(S1114)。

【0096】次に、行・文字抽出部207が局所画像メモリ204に記憶された文字認識実行領域の画像データから文字および行の画像データを抽出する(S1115)。そして、行・文字抽出部207が抽出した文字を用いて、文字認識部208が文字認識を行う(S1116)。

【0097】その後、オペレータによって指定された指定領域全体についての文字認識が終了したか否かを判断する(S1117)。そして、指定領域に文字認識されていない部分があると判断された場合には、撮像デバイス制御部206がCCDカメラ102の撮像位置をシフトさせ(S1118)、次の撮像を行った後(S1113)、ステップS1113～S1118を繰り返す。この動作は、ステップS1117において指定領域全体についての文字認識が終了したと判断されるまで繰り返される。

【0098】ステップS1117で、指定領域全体についての文字認識が終了したと判断された場合には、文字認識部208の認識結果が認識結果格納・出力部209に送られる(S1119)。これにより、文字認識部208によって認識された文字がディスプレイ装置(図示せず)に表示される等して、情報処理装置による処理が終了する。

【0099】なお、文字認識領域の2箇所以上を指定できるようにしたい場合には、ステップS1103～S1115を繰り返し実行するように、画像処理装置を構成すればよい。

【0100】また、文字認識実行領域の指定が正常に終了したことを、ピープ音やランプ等によって、オペレータに告知することとしてもよい。

【0101】このように、この実施の形態に係る画像処理装置によれば、原稿上で始点と終点とを指し示すことによって文字認識を行いたい領域を指定することができるので、文字認識を行いたい領域をオペレータが任意に指定することができ、従って、装置の使い勝手を向上させることができる。

【0102】また、処理速度を向上させることができる点、メモリ容量を低減することができる点、撮像素子(ここではCCDカメラ102)として比較的低解像度のもので使用することができる点、設置スペースを小さくすることができる点および本や雑誌等をオペレータが原稿を押さえる必要や綴じ合わせ部分を破損するおそれがない点は、第1の実施の形態と同様である。

【0103】第4の実施の形態

次に、この発明の第4の実施の形態について、図14～図16を用いて説明する。

【0104】ここで、この実施の形態では、図1(B)

に示した場合と同様、撮像デバイスとしてのCCDカメラ102は、オペレータが手に持って操作するように構成されている。

【0105】また、後述するように、本体部101は、オペレータがCCDカメラ102を原稿に近づけることによって指定した領域についてのみ文字認識を行うように構成されている。

【0106】図14は、この実施の形態の情報処理装置に係る回路構成の要部を示すブロック図である。

【0107】同図において、図2と同じ符号を付した構成部は、それぞれ図2の場合と同じものを示している。

【0108】また、スナップショットメモリ1401は、後述するようにしてCCDカメラ102が所定時間ごとに原稿を撮像することによって得られた画像データを、順次入力して記憶する。

【0109】認識領域識別部1402は、スナップショットメモリ1401から画像データを取り込んで、文字認識を行う領域を判別する。

【0110】次に、この実施の形態に係る情報処理装置の動作について、図15および図16を用いて説明する。

【0111】オペレータは、CCDカメラ102を手に持ち、このCCDカメラ102によって原稿全体を撮像させる。

【0112】このとき、撮像デバイス入力部201は、CCDカメラ102から画像データを取り込み(S1501)、スナップショットメモリメモリ1401に格納する(S1502)。

【0113】さらに、認識領域識別部1402が、この画像データをスナップショットメモリ1401から読み出し、文字認識実行領域(図16(A)参照)を抽出する(S1503)。この実施の形態では、所定長さ以上の白ランで囲まれた領域を、文字認識実行領域であると判断する。これにより、表領域等は文字認識実行領域に含まれるが、小さい黒点のごときは含まれないこととなる。なお、白ランの長さのしきい値は、設計段階で任意に定めることができる。

【0114】続いて、認識領域識別部1402が、文字認識実行領域が抽出されたか否かを判断し(S1504)、抽出されなければステップS1501～S1504を繰り返す。

【0115】一方、ステップS1504で文字認識実行領域が抽出された場合は、その領域の数が1個か否かを判断する(S1505)。そして、文字認識実行領域が2個以上である場合は、ステップS1501～S1505を繰り返す。

【0116】オペレータがCCDカメラ102を原稿に徐々に近づけていくと、図16の(A)～(E)に示したように、原稿を徐々に拡大した画像データが撮像デバイス201に取り込まれる。そして、この画像データの

うち、所定長さ以上の白ランで囲まれた領域が1個になると(図16(C)のG<sub>4n</sub>参照)、ステップS1505で、文字認識実行領域が1個であると判断される。この場合には、このときの画像データが、行・文字抽出部207に送られる。

【0117】行・文字抽出部207は、この画像データから、文字および行の画像データを抽出する(S1506)。そして、行・文字抽出部207が抽出した文字を用いて、文字認識部208が文字認識を行う(S1507)。

【0118】このような動作を繰り返すことにより、オペレータが望む領域の文字認識を順次行うことができる。

【0119】また、文字認識を行う領域が確定した時点(ステップS1505で所定長さ以上の白ランで囲まれた領域が1個になった時点)で、ピープ音やランプ等によるオペレータへの告知を行うこととしてもよい。

【0120】このように、この実施の形態に係る画像処理装置によれば、文字認識を行いたい領域をオペレータが指定する際に、CCDカメラ102を原稿に近づけるだけでよいので、従来のようにディスプレイ装置の画面上で指定する場合と比較して、装置の使い勝手を向上させることができる。

【0121】また、処理速度を向上させることができる点、メモリ容量を低減することができる点、撮像素子(ここではCCDカメラ102)として比較的低解像度のものを使用することができる点、設置スペースを小さくすることができる点および本や雑誌等をオペレータが原稿を押さえる必要や綴じ合わせ部分を破損するおそれがない点は、第1の実施の形態と同様である。

【0122】

【発明の効果】以上詳細に説明したように、この発明によれば、省スペース型で、読み取り時間が速く、オペレータの作業負担も小さい情報処理装置を提供することができる。

【図面の簡単な説明】

【図1】第1の実施の形態に係る情報処理装置の全体構成を概念的に示す外観図である。

【図2】第1の実施の形態の情報処理装置に係る回路構成の要部を示すブロック図である。

【図3】第1の実施の形態に係る情報処理装置の動作を

説明するための概略フローチャートである。

【図4】第1の実施の形態に係る情報処理装置の動作を説明するための概念図である。

【図5】第2の実施の形態に係る情報処理装置の全体構成を概念的に示す外観図である。

【図6】第2の実施の形態の情報処理装置に係る回路構成の要部を示すブロック図である。

【図7】第2の実施の形態に係る情報処理装置の動作を説明するための概略フローチャートである。

10 【図8】第2の実施の形態に係る情報処理装置の動作を説明するための概略フローチャートである。

【図9】第2の実施の形態に係る情報処理装置の動作を説明するための概念図である。

【図10】第3の実施の形態の情報処理装置に係る回路構成の要部を示すブロック図である。

【図11】第3の実施の形態に係る情報処理装置の動作を説明するための概略フローチャートである。

【図12】第3の実施の形態に係る情報処理装置の動作を説明するための概略フローチャートである。

20 【図13】第3の実施の形態に係る情報処理装置の動作を説明するための概念図である。

【図14】第4の実施の形態の情報処理装置に係る回路構成の要部を示すブロック図である。

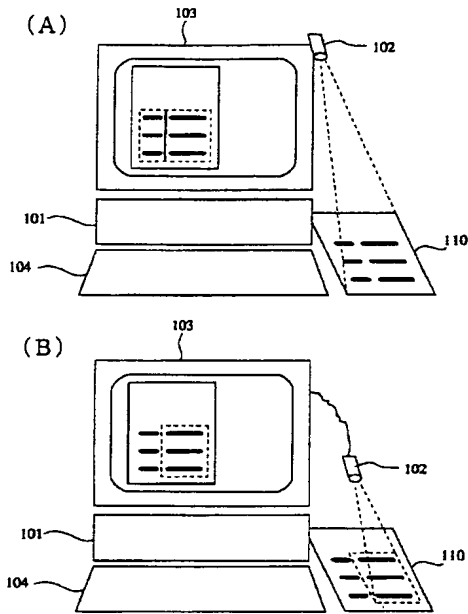
【図15】第4の実施の形態に係る情報処理装置の動作を説明するための概略フローチャートである。

【図16】第4の実施の形態に係る情報処理装置の動作を説明するための概念図である。

【符号の説明】

- 101 本体部
- 102 CCDカメラ
- 103 ディスプレイ装置
- 104 キーボード
- 201 撮像デバイス入力部
- 202 広域画像メモリ
- 203 局所画像メモリ
- 204 前処理・認識領域識別部
- 205 座標データ記憶部
- 206 撮像デバイス制御部
- 207 行・文字抽出部
- 208 文字認識部
- 209 認識結果格納・出力部

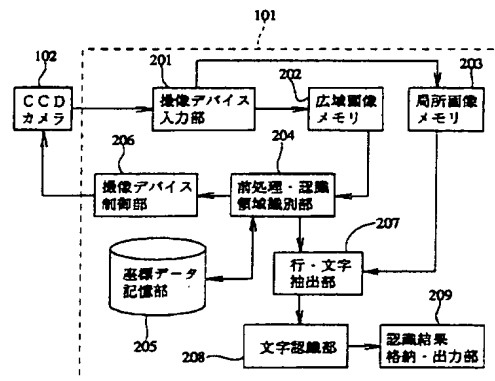
【図1】



101: 本体部  
102: CCDカメラ  
103: ディスプレイ装置  
104: キーボード  
110: 原稿

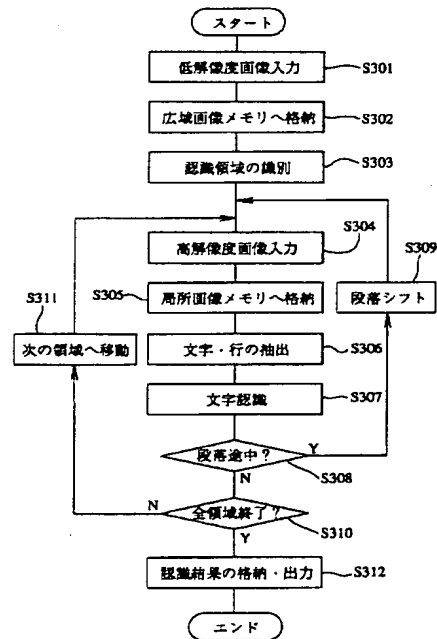
第1の実施の形態の外観図

【図2】

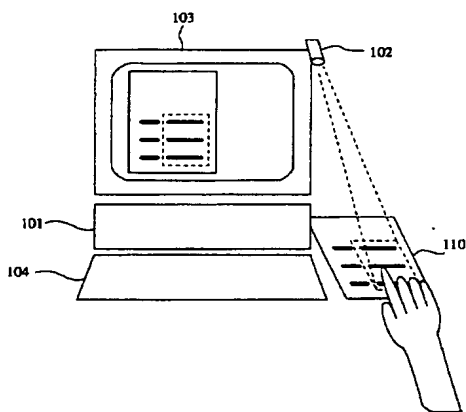


第1の実施の形態の構成図

【図3】

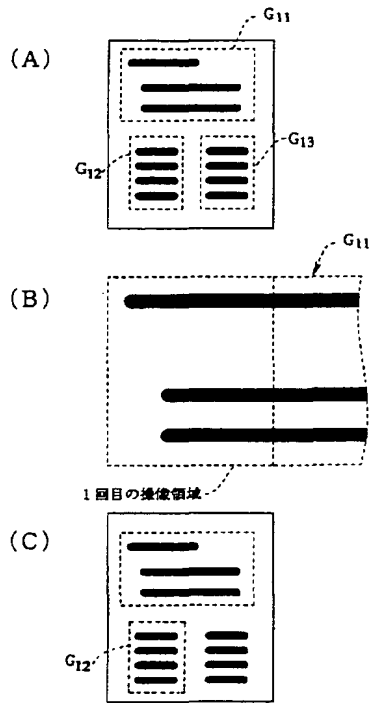


【図5】



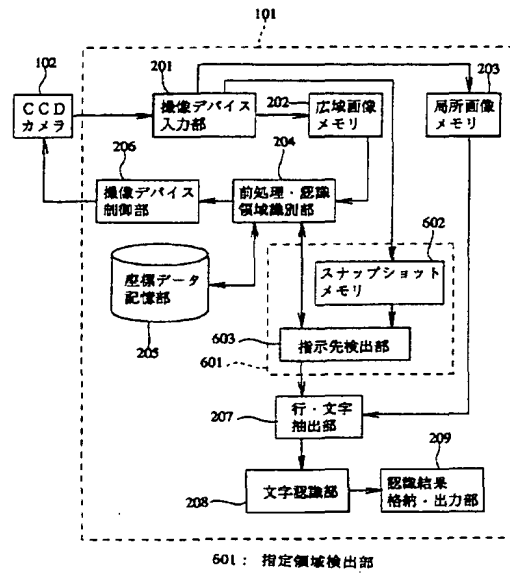
第2の実施の形態の外観図

【図4】



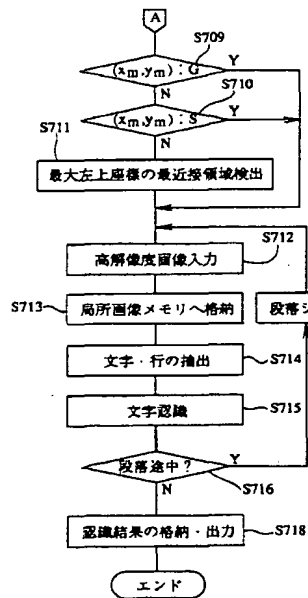
第1の実施の形態の動作説明図

【図6】

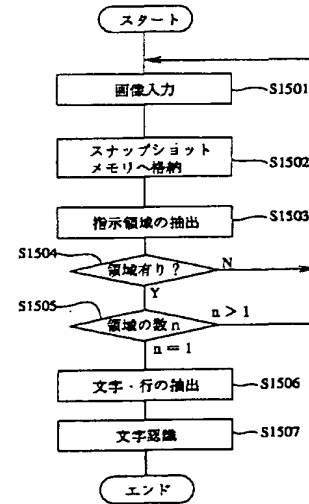


第2の実施の形態の構成図

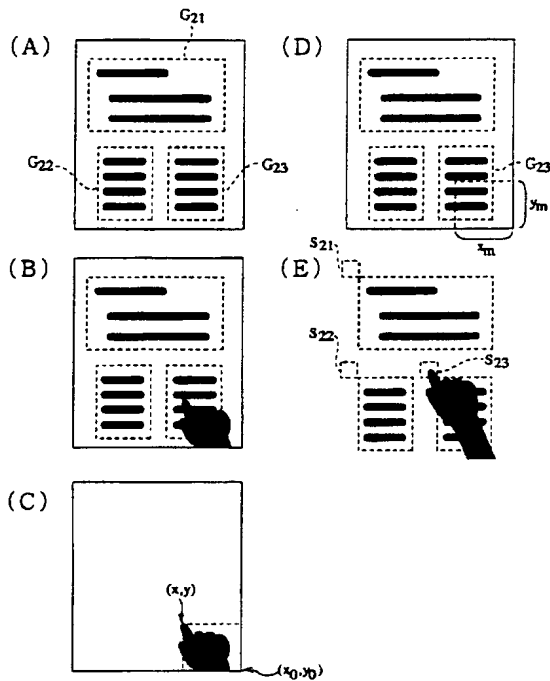
【図8】



【図15】

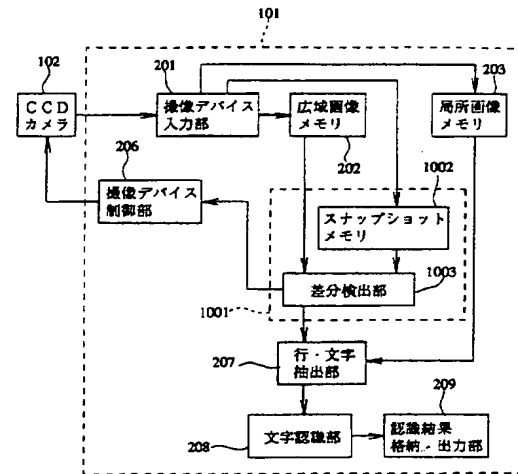


【図9】



第2の実施の形態の動作説明図

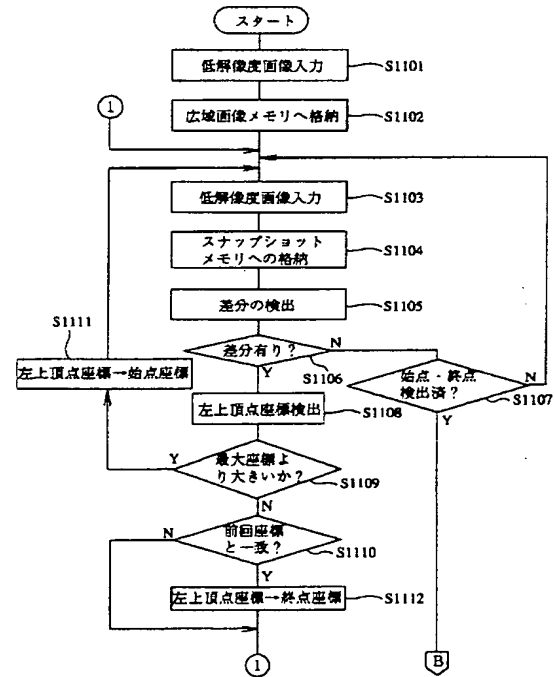
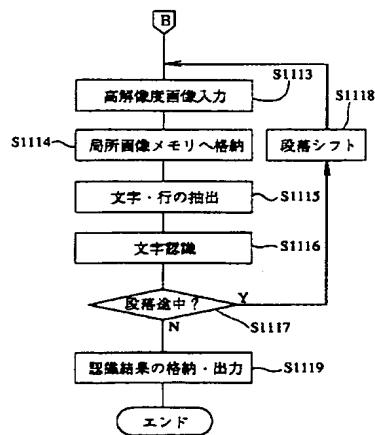
【図10】



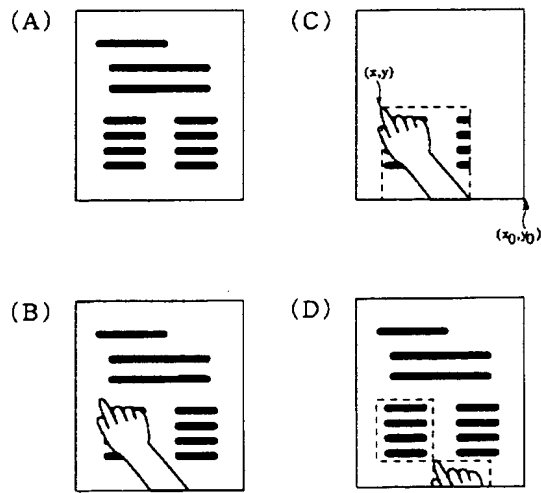
第3の実施の形態の構成図

【図11】

【図12】

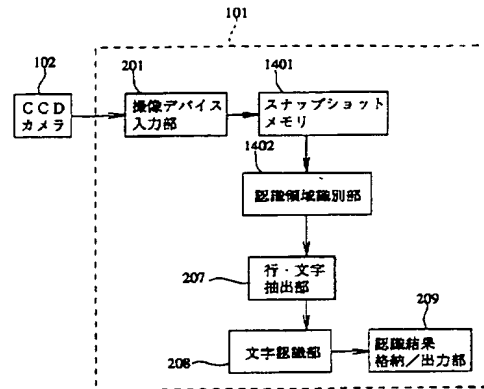


【図 13】



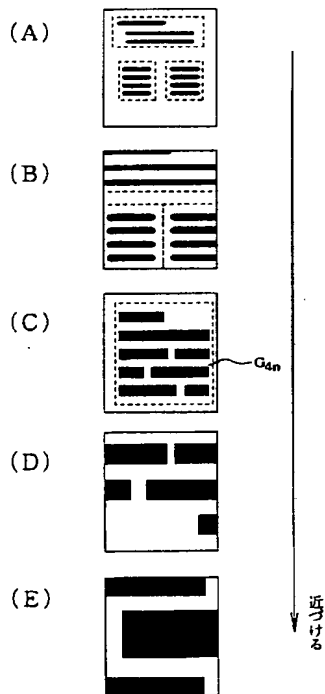
第3の実施の形態の動作説明図

【図 14】



第4の実施の形態の構成図

【図 16】



第4の実施の形態の動作説明図